DOCUMENT RESUME

ED 234 907 PS 013 847

AUTHOR Adolf, Jane W.
TITLE An Examination of Stallings' "Implementation and Child Effects of Teaching Practices in Follow Through Classrooms."
PUB DATE Sep 83
NOTE 12p.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Early Childhood Education; *Economically Disadvantaged; *Evaluation Methods; *Outcomes of Education; *Program Evaluation; Teacher Evaluation; *Teaching Models
IDENTIFIERS Evaluation Reports; *Evaluation Utilization; *Project Follow Through

ABSTRACT

The purpose of this paper is to critique Stallings' 1975 evaluation of Project Follow Through (PFT) in order to examine methodological issues inherent in evaluation research. Following a summary of the objectives, methods, and conclusions of Stallings' study, the critique focuses on four elements of research evaluation: internal validity, external validity, reliability, and relevance. Criticism of the PFT evaluation includes the following assertions: (1) Internal validity may be affected by a discrepancy in the number of variables used to describe program models and the number used to assess individual programs; (2) External validity may be questionable due to nonrandom assignment of the sample to treatment conditions; and (3) Reliability may be reduced as a result of confounding observers with program models, possible inconsistency of observations, and lack of control over extraneous variables. The relevance of Stallings' evaluation is perceived as being substantial in that it produced valuable implications for the development of educational programs and contributed methodological innovations for studying large educational projects. Concluding remarks review major findings of the PFT evaluation and, emphasizing the utility of the Classroom Observation Instrument, point out implications of Stallings' work for future evaluation research. (RH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☒ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED234907

AN EXAMINATION OF STALLINGS' "IMPLEMENTATION AND CHILD EFFECTS OF TEACHING PRACTICES IN FOLLOW THROUGH CLASSROOMS"

Jane W. Adolf
University of Minnesota
Institute of Child Development
September, 1983

PS 013847

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Jane W. Adolf
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

An Examination of Stallings'

"Implementation and Child Effects of Teaching Practices In

Follow Through Classrooms"

Project Follow Through was established by the U.S. Congress in 1967 when it became apparent that a program was needed in the early grades of public school that was articulated with Project Head Start goals and approaches and, therefore, would provide a comparable educational program for economically deprived children over a longer period of time (Stallings, 1975). A major goal of the project was to examine the differential effectiveness of programs based on divergent educational and developmental theories. Twenty-two Follow Through sponsors implemented models, each with a different approach to improving the educational achievement of economically disadvantaged children. Now, over ten years later, evaluations of Project Follow Through programs have accrued in attempts to determine the effects of this project, and to offer implications for programmatic change.

The purpose of this paper is to critique one evaluation of Project Follow Through in order to examine methodological issues inherent in evaluative research. The evaluation under investigation is Stalling's 1975 report titled "Implementation and Child Effects of Teaching Practices in Follow Through Classrooms". Before examining salient strengths and weaknesses of Stalling's evaluation report, however, a summary of the objectives, methods and conclusions of her study is offered.

In most evaluations of Project Follow Through, the major emphasis has been to determine child outcomes in terms of child performance (Stallings, 1975). Due to the number of different Follow Through models, however, it is apparent that without identification of variables in the classroom affecting performance, the

measure of child outcomes means very little.

In Stalling's evaluation she examined two pervasive questions with respect to Project Follow Through: (1) Are the observed classroom teaching practices for Follow Through participants consistent with the goals of the sponsor's model, and (2) How are teaching practices related to children's outcomes? Stalling's study focuses upon whether sponsors can deliver their educational models to diverse kinds of communities. An educational theory can only be proved to be effective, as measured by child outcomes, if the teachers and aides carry out program specifications.

In order to determine whether the sponsors were effective in getting teachers to practice prescribed methods in the classroom, it was necessary to observe the classrooms systematically. The Classroom Observation Instrument (COI) was designed to record classroom arrangements and elements of events considered educationally significant by the Follow Through sponsors. The COI included over 600 behaviors that teachers and students exhibit. By recording actual classroom practices with their prescribed models, a comparison between the two entities can be undertaken, and any discrepancies identified.

To determine the relationship between teaching practices and children's outcomes, a number of dependent measures were collected by the evaluators: Metropolital Achievement Test, the Raven's Coloured Progressive Matrices, Intellectual Achievement Responsibility Scale, absence rate, and desired child behavior. It was through relating child outcomes to instructional processes, regardless of where these processes occurred, that antecedents to these outcomes could be investigated.

The vast majority of classrooms of the seven Follow Through sponsors investigated were found to be significantly different from the comparison classrooms (Stallings, 1975). Further, the different classrooms of a given sponsor were observed to be remarkably similar to each other regardless of where they were

being implemented. In most cases, programs reflected the goals and model of their particular sponsor. On the basis of observation, most classrooms could be correctly classified as to sponsor without prior knowledge of what model the teacher was attempting to put into practice.

The study of instructional procedures used in classrooms and the achievement of children indicates that time spent in reading and math activities and a high rate of drill, practice and praise contribute to higher reading and math scores (Stallings, 1975).

Based on these findings, the evaluators concluded that what occurs within a classroom is related to achievement in basic skills, good attendance and desired child behaviors (Stallings, 1975). Further, the seven Follow Through models considered in the study bring different strengths to their pupils and each is bringing advantages not usually found in traditional classrooms.

Before examining specific aspects of Stallings' work, it is worthwhile to identify the criteria upon which the study will be critiqued. A review of evaluation literature provides an insight to the many opinions that exist concerning the manner in which evaluation should be undertaken (Hamilton, et al, 1977, Scriven, 1973, Suchman, 1967). In order to narrow this examination, the criteria for a good evaluation as enumerated by Stufflebeam et al (1971) will be the basis for the discussion.

According to Stufflebeam, evaluation may be defined as the process of delineating, obtaining or providing information necessary to serve decision-making (Guba, 1975). He proposes the following list as critical elements in a research evaluation: internal validity, external validity, reliability, objectivity, relevance, importance, scope, credibility, timeliness, pervasiveness, and efficiency. In the present examination of Stallings' work the discussion will focus on the following attributes: internal validity, external validity, reliability, and relevance. In addition, an examination of the significance of

Stallings' study will be undertaken to highlight the contributions that her evaluation has made in terms of Project Follow Through and evaluation methodology in general. These elements have been selected because they represent major features of any evaluation and offer interesting discussion with respect to Stallings' work.

The first aspect of Stallings' study to be examined is its internal validity. Internal validity asks the question: Did the experimental manipulation really make a significant difference (Kerlinger, 1973)? According to Stufflebeam et al (1971), in an evaluation this criterion implies a one-to-one correspondence between the evaluation information and the phenomena which it purports to describe.

When reviewing the objectives of Stallings' evaluation, the methods used to obtain the objectives and the study's conclusion, the efforts of the evaluators to measure significant variables in a well-planned manner are quite impressive. The development of the COI was a major contributor to really assessing the continuity of like-models across varying settings and the existence of differentiation between models. It appears that Stallings has identified a significant variable and examined it in a thorough and systematic manner. Indeed, the information attained concerning program consistency is an important control feature for conclusions about the effects of programs as measured by child outcomes. One problem with internal validity, however, is that the COI and its measure of classroom practices is questionable because sponsors were given a much smaller number of variables to describe their models than were used to assess individual programs. This implies that Stallings collected more data than necessary and complicated her analysis and interpretations.

The internal validity concerning child outcomes is only as good as the measurements that were used for assessment. Stallings attained a number of measures from standard tests and also from information from the COI. The tests

represent a spectrum of effects that children may experience in Follow Through programs, and Stallings' efforts to get an idea of the whole child are commendable. However, one must also consider the validity of the measures utilized in her study. The reader must be aware of test limitations, and use caution when making conclusions based on their results.

The second criterion of a good evaluation is its external validity. This aspect deals with the generalizability of findings beyond the sample (Guba, 1975). Questions of sample representativeness, similarity of conditions to the population, reactive effects and other factors are likely to have serious bearing on the generalizability of results. Because Project Follow Through is a large scale project desiring widespread effects on the target population, the ability to generalize from the study's sample is important.

Selection and observation os sites was based upon the following criteria: (1) that they were among the sites where pupil testing was to occur in Spring, 1973 as part of the Follow Through evaluation; (2) that each sponsor would, as much as possible, have a balanced geographic distribution of sites, which included urban-rural and north-south projects; and (3) that each sponsor would have included at least two sites which he considered well implemented.

It is interesting to note that one of the major objectives of Stallings' evaluation was to study the generalizability of child outcomes according to program. She hoped to determine whether sponsors' models were being implemented uniformly so that generalizations made from these models would be more reliable. Her sample is representative of the population involved with Project Follow Through. However, one must ask the question, how representative are the subjects in the project a reflection of the entire economically deprived population? Limitations on generalizability exist if those involved with Project Follow Through are significantly different than the larger population. Stallings' sam-

ple was essentially an accidental one; random assignment according to treatment was not possible. In any case, Project Follow Through findings should be limited to the lower SES population.

The question of reliability asks whether the information gained from an evaluation is consistent (Guba, 1975). In large scale evaluations, controlling confounding variables is not an easy task. This problem is especially prevalent in educational research conducted in a natural environment. However, all efforts should be made to minimize extraneous variability and carefully describe any factors that may influence the results.

A few of the reliability problems are presented to highlight the problems that emerge with respect to controlling variables with a large sample in an educational setting. A problem exists with respect to reliability of observations. A major problem with this study was a confounding of observers with models. Not only did observers make observations at only one site, but previous knowledge of the program might have introduced biases. Stallings checks on the reliability of observations were essentially checks on the reliability of observers' skills rather than the reliability of the coding system. The consistency of observations is also called into question when considering the variability among teachers and children not only between models, but also within individuals on a day-to-day basis. A third consideration should be made concerning certain extraneous variables that may effect results. The evaluators could not control for subject attrition of differing amounts of program exposure except through statistical manipulations.

The most interesting aspect to critiquing an evaluation is examining its implications and significance. Stallings' report is particularly fascinating because it not only offers substantial implications for the development of educational programs, but also offers significant contributions to methodological innovations for studying large, educational projects.

The major findings of this study are reviewed so implications from the results may be discussed (Stallings, 1975):

(1) Seven sponsors' educational models each had unique features that were consistently present in all five of a sponsor's sites. The seven sponsor's models are considered to be well implemented and could be expected to be implemented in other locations similar to those in the present study.

(2) Classroom instructional practices in first- and third-grade classrooms have been identified which contribute to higher test scores, other desired child behavior, and lower absence rates.

(3) Highly controlled classroom environments in which teachers used systematic instruction and a high rate of positive reinforcement contributed to higher scores in mathematics and reading. Flexible classroom environments which provided more exploratory materials and allowed for more choice on the part of the child contributed to higher scores on a test of nonverbal reasoning, lower absence rates, and a willingness of the part of the children to work independently.

Stallings has shown that the classroom behavior of both teachers and students are not only different according to sponsorship but that the type of sponsor is relevant to the kinds of outcomes produced by the child. These findings are significant because they offer decision makers an interesting ethical problem. Given the information that program processes effect child outcomes, a decision must be made as to what are the desired outcomes for children from economically deficient backgrounds. (Should policy makers prescribe the direction of educational and individual development? It is rather ironic that large amounts of Federal funds are poured into programs that have differential effects on its participants. Gains in specific academic areas and even approaches to learning can be attributed to specific elements in programs. Therefore children are

developing differently according to the project with which they are associated. Who should decide how children's development should be supplemented? In addition, who should decide which children should participate in these programs? Is it really fair, given that special programs affect child outcomes, that all children do not receive the very best education?)

The answers to such questions must includemore than the ethical issues. The financing of special projects, their development, implementation and supervision, is an expensive endeavor. (How much are we willing to spend to give all children the most beneficial, ethically sound education? If a cost-effectiveness evaluation were performed on the programs and their subsequent effects on children, how do we determine if it is all worth it? If we are trying to raise people from low socio-economic status, perhaps potential participants would rather receive a monetary compensation versus expenditures on their behalf to the educational system.)

A critique of Stallings' work would not be complete without a discussion of its implications for future research in evaluation. The methodology of Stalling's study is undoubtedly unfamiliar to many psychologists who are used to carefully controlled experimental designs that explore the relationships among a small set of variables (Resnick and Leinhardt, 1975). When one undertakes educational evaluation, especially on a large scale, the number of uncontrollable variable presents major methodological problems. It is for this reason that some educational evaluators feel that educational research has focused on controllable situations that offer answers to trivial questions (Atkins, 1977). Stallings' efforts to examine a multitude of complex variables in a number of diverse settings provides needed insight to overall programmatic effects. She has identified important variables, and her examination of them are admirable and innovative.

One of Stallings' major contributions to future research is the introduction of the Classroom Observation Instrument. This detailed coding system has not only increased knowledge of classroom practices, but appears to be adaptable for use in answering other educational questions. One cannot help but infer that one of the major unspoken objectives of Stallings' research was to develop an observational instrument that might provide dependent measures of classroom practices for use in future studies.

To detect the real effects of programmatic innovations in education it seems logical to utilize both measures of classroom practices and child outcomes. One measurement without the other does not provide conclusive evidence concerning the success or failure of a program. Be defining the unit of analysis at the classroom versus program model level, a clearer impression of what is actually occurring can be identified. Stallings' monograph offers a significant step toward integrating all of these contributing factors to the assessment of programmatic implementation and effects. As such, "Implementation and Child Effects of Teaching Practices in Follow Through Classrooms" is an important monograph offering indepth reporting, extensive analysis, and more conclusive findings.

REFERENCES

Atkins, M. Evaluation for content improvement in Beyond the Numbers Game, Hamilton et al (Eds.) Berkeley: McCutchan, 1977.

Guba, Egon. Problems in Utilizing the Results of Evaluation. Journal of Research and Development in Education, 1975, Volume 8, No. 3, pp. 42-54.

Hamilton, D., MacDonald, D., King, C., Jenkins, D. & Parlett, M. (Eds.) Beyond the Numbers Game. Berkeley: McCutchan Publishing Corp., 1977.

Kerlinger, F. Foundations of Behavioral Research, 2nd Ed. New York: Holt, Rinehart and Winston, 1973.

Resnick, L. & Leinhardt, G. Commentary in Implementation and Child Effects of Teaching Practices in Project Follow Through Classrooms, Monographs of the Society for Research in Child Development, 1975, 40, (7-8, Serial No. 163).

Scriven, M. Goal-free Evaluation. E. R. House (Ed.) School Evaluation: The Politics and Processes. Berkeley: McCutchan, 1973, 48-67.

Stallings, J. Implementation and Child Effects of Teaching Practices in Follow Through Classrooms. With Commentary by Millie Almy, and by Lauren B. Resnick and Gaea Leinhardt. Monographs of the Society for Research in Child Development, 1975, 40, (7-8, Serial No. 163).

Stufflebeam, D. L. et al. Educational Evaluation and Decision-Making. Itasca, IL.: Peacock, 1971.

Suchman, E. A. Evaluative Research. New York: Sage, 1967.